



日本特許庁
JAPAN PATENT OFFICE

CF015501 US/mas
H.S.

09/091,389
Hiroyuki Kimura
July 27, 2001

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日
Date of Application:

2000年 7月 6日

出願番号
Application Number:

特願2000-204959

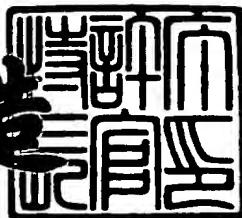
出願人
Applicant(s):

キヤノン株式会社

2001年 7月 19日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2001-3064555

【書類名】 特許願

【整理番号】 4176016

【提出日】 平成12年 7月 6日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明の名称】 音声合成装置、音声合成システム、音声合成方法及び記憶媒体

【請求項の数】 50

【発明者】

【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

【氏名】 木村 裕行

【発明者】

【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

【氏名】 後藤 裕典

【特許出願人】

【識別番号】 000001007

【氏名又は名称】 キヤノン株式会社

【代表者】 御手洗 富士夫

【代理人】

【識別番号】 100081880

【弁理士】

【氏名又は名称】 渡部 敏彦

【電話番号】 03(3580)8464

【手数料の表示】

【予納台帳番号】 007065

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9703713
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成装置、音声合成システム、音声合成方法及び記憶媒体

【特許請求の範囲】

【請求項1】 テキストデータを合成音声に変換して出力する音声合成装置であって、

前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力する音声出力手段とを有することを特徴とする音声合成装置。

【請求項2】 前記音声出力手段は、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力し、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力することを特徴とする請求項1記載の音声合成装置。

【請求項3】 前記音声出力手段は、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする請求項1又は2記載の音声合成装置。

【請求項4】 テキストデータを合成音声に変換して出力する音声合成装置であって、

前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設ける音声出力手段とを有することを特徴とする音声合成装置。

【請求項5】 前記空白時間は、任意に設定可能であることを特徴とする請求項4記載の音声合成装置。

【請求項6】 テキストデータを合成音声に変換して出力する音声合成装置であって、

前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の

音声波形の再生開始前に、予め用意された特定の音声合成波形を再生する音声出力手段とを有することを特徴とする音声合成装置。

【請求項7】 前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする請求項6記載の音声合成装置。

【請求項8】 遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話を行うシステムに適用可能であることを特徴とする請求項1乃至7の何れかに記載の音声合成装置。

【請求項9】 テキストデータを合成音声に変換して出力する音声合成装置と、該音声合成装置に前記テキストデータを送信する外部装置とを具備してなる音声合成システムであって、

前記音声合成装置は、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力する音声出力手段とを有することを特徴とする音声合成システム。

【請求項10】 前記音声合成装置の前記音声出力手段は、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力し、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力することを特徴とする請求項9記載の音声合成システム。

【請求項11】 前記音声合成装置の前記音声出力手段は、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする請求項9又は10記載の音声合成システム。

【請求項12】 テキストデータを合成音声に変換して出力する音声合成装置と、該音声合成装置に前記テキストデータを送信する外部装置とを具備してなる音声合成システムであって、

前記音声合成装置は、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設け

る音声出力手段とを有することを特徴とする音声合成システム。

【請求項13】 前記空白時間は、任意に設定可能であることを特徴とする請求項12記載の音声合成システム。

【請求項14】 テキストデータを合成音声に変換して出力する音声合成装置と、該音声合成装置に前記テキストデータを送信する外部装置とを具備してなる音声合成システムであって、

前記音声合成装置は、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生する音声出力手段とを有することを特徴とする音声合成システム。

【請求項15】 前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする請求項14記載の音声合成システム。

【請求項16】 遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話を行うシステムに適用可能であることを特徴とする請求項9乃至15の何れかに記載の音声合成システム。

【請求項17】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、

前記テキストデータの音声波形を生成する音声波形生成工程と、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力する音声出力工程とを有することを特徴とする音声合成方法。

【請求項18】 前記音声出力工程では、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力し、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力することを特徴とする請求項17記載の音声合成方法。

【請求項19】 前記音声出力工程では、音声再生待ちの音声波形の数に応

じた細かいステップで再生スピードを上げることが可能であることを特徴とする請求項17又は18記載の音声合成方法。

【請求項20】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、

前記テキストデータの音声波形を生成する音声波形生成工程と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設ける音声出力工程とを有することを特徴とする音声合成方法。

【請求項21】 前記空白時間は、任意に設定可能であることを特徴とする請求項20記載の音声合成方法。

【請求項22】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、

前記テキストデータの音声波形を生成する音声波形生成工程と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生する音声出力工程とを有することを特徴とする音声合成方法。

【請求項23】 前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする請求項22記載の音声合成方法。

【請求項24】 遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話を行うシステムに適用可能であることを特徴とする請求項17乃至23の何れかに記載の音声合成方法。

【請求項25】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、

前記音声合成方法は、前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力するよ

うに制御する音声出力ステップとを有することを特徴とする記憶媒体。

【請求項26】 前記音声出力ステップでは、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力するように制御し、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力するよう制御することを特徴とする請求項25記載の記憶媒体。

【請求項27】 前記音声出力ステップでは、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする請求項25又は26記載の記憶媒体。

【請求項28】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、

前記音声合成方法は、前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設けるように制御する音声出力ステップとを有することを特徴とする記憶媒体。

【請求項29】 前記空白時間は、任意に設定可能であることを特徴とする請求項28記載の記憶媒体。

【請求項30】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、

前記音声合成方法は、前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生するように制御する音声出力ステップとを有することを特徴とする記憶媒体。

【請求項31】 前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形で

あることを特徴とする請求項30記載の記憶媒体。

【請求項32】 遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話を行うシステムに適用可能であることを特徴とする請求項25乃至31の何れかに記載の記憶媒体。

【請求項33】 テキストデータを合成音声に変換して出力する音声合成装置であって、

前記テキストデータを入力する入力手段と、

前記テキストデータの音声波形を生成する音声波形生成手段と、

前記音声波形に係る音声を出力する音声出力手段と、

第1の前記テキストデータに係る音声の出力中に第2の前記テキストデータの入力による音声波形を検出した場合に、前記第1のテキストデータに係る音声の出力を終了した後前記第2のテキストデータに係る音声を出力するよう前記音声出力手段を制御する制御手段とを備えることを特徴とする音声合成装置。

【請求項34】 前記制御手段は前記第2のテキストデータによる音声波形の検出に応じて前記第1のテキストデータに係る音声波形の再生速度を通常速度よりも速くするよう前記音声出力手段を制御することを特徴とする請求項33記載の音声合成装置。

【請求項35】 前記制御手段は前記第1のテキストデータに係る音声出力の終了後、所定期間経過した後に前記第2のテキストデータに係る音声の出力を開始するよう前記音声出力手段を制御することを特徴とする請求項33記載の音声合成装置。

【請求項36】 前記制御手段は前記第1のテキストデータに係る音声出力の終了後、所定の音声を出力した後に前記第2のテキストデータに係る音声を出力するよう前記音声出力手段を制御することを特徴とする請求項33記載の音声合成装置。

【請求項37】 前記制御手段は前記第1のテキストデータに係る音声と前記第2のテキストデータに係る音声とを通常の再生速度で出力することを特徴とする請求項33記載の音声合成装置。

【請求項38】 前記音声波形生成手段により生成された音声波形データを記憶する記憶手段を備え、前記制御手段は前記記憶手段に記憶されている前記入力テキストデータに応じた音声波形データの数に応じて前記音声波形の再生速度を変更するよう前記音声出力手段を制御することを特徴とする請求項33記載の音声合成装置。

【請求項39】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、

前記テキストデータを入力する入力工程と、

前記テキストデータの音声波形を生成する音声波形生成工程と、

前記音声波形に係る音声を出力する音声出力工程と、

第1の前記テキストデータに係る音声の出力中に第2の前記テキストデータの入力による音声波形を検出した場合に、前記第1のテキストデータに係る音声の出力を終了した後前記第2のテキストデータに係る音声を出力するよう前記音声出力工程を制御する制御工程とを有することを特徴とする音声合成方法。

【請求項40】 前記制御工程では前記第2のテキストデータによる音声波形の検出に応じて前記第1のテキストデータに係る音声波形の再生速度を通常速度よりも速くするよう前記音声出力工程を制御することを特徴とする請求項39記載の音声合成方法。

【請求項41】 前記制御工程では前記第1のテキストデータに係る音声出力の終了後、所定期間経過した後に前記第2のテキストデータに係る音声の出力を開始するよう前記音声出力工程を制御することを特徴とする請求項39記載の音声合成方法。

【請求項42】 前記制御工程では前記第1のテキストデータに係る音声出力の終了後、所定の音声を出力した後に前記第2のテキストデータに係る音声を出力するよう前記音声出力工程を制御することを特徴とする請求項39記載の音声合成方法。

【請求項43】 前記制御工程では前記第1のテキストデータに係る音声と前記第2のテキストデータに係る音声とを通常の再生速度で出力することを特徴とする請求項39記載の音声合成方法。

【請求項4 4】 前記音声波形生成工程により生成された音声波形データを記憶する記憶工程を有し、前記制御工程では前記記憶工程に記憶されている前記入力テキストデータに応じた音声波形データの数に応じて前記音声波形の再生速度を変更するよう前記音声出力工程を制御することを特徴とする請求項3 9記載の音声合成方法。

【請求項4 5】 テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、

前記音声合成方法は、

前記テキストデータを入力する入力ステップと、

前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、

前記音声波形に係る音声を出力するように制御する音声出力ステップと、

第1の前記テキストデータに係る音声の出力中に第2の前記テキストデータの入力による音声波形を検出した場合に、前記第1のテキストデータに係る音声の出力を終了した後前記第2のテキストデータに係る音声を出力するよう前記音声出力ステップを制御する制御ステップとを有することを特徴とする記憶媒体。

【請求項4 6】 前記制御ステップでは前記第2のテキストデータによる音声波形の検出に応じて前記第1のテキストデータに係る音声波形の再生速度を通常速度よりも速くするよう前記音声出力ステップを制御することを特徴とする請求項4 5記載の記憶媒体。

【請求項4 7】 前記制御ステップでは前記第1のテキストデータに係る音声出力の終了後、所定期間経過した後に前記第2のテキストデータに係る音声の出力を開始するよう前記音声出力ステップを制御することを特徴とする請求項4 5記載の記憶媒体。

【請求項4 8】 前記制御ステップでは前記第1のテキストデータに係る音声出力の終了後、所定の音声を出力した後に前記第2のテキストデータに係る音声を出力するよう前記音声出力ステップを制御することを特徴とする請求項4 5記載の記憶媒体。

【請求項49】 前記制御ステップでは前記第1のテキストデータに係る音声と前記第2のテキストデータに係る音声とを通常の再生速度で出力するように制御することを特徴とする請求項45記載の記憶媒体。

【請求項50】 前記音声波形生成ステップにより生成された音声波形データを記憶するように制御する記憶ステップを有し、前記制御ステップでは前記記憶ステップに記憶されている前記入力テキストデータに応じた音声波形データの数に応じて前記音声波形の再生速度を変更するよう前記音声出力ステップを制御することを特徴とする請求項45記載の記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声合成装置、音声合成システム、音声合成方法及び記憶媒体に関し、特に、テキストデータを合成音声に変換して出力する場合に好適な音声合成装置、音声合成システム、音声合成方法及び記憶媒体に関する。

【0002】

【従来の技術】

従来より、文字情報を音声出力する機能を有する音声合成装置がある。従来の音声合成装置において、音声出力するデータは、予め電子化されたテキストデータとして準備されたものでなければならなかった。即ち、そのテキストデータは、パーソナルコンピュータ上におけるエディタやワードプロセッサなどで作成した文章や、インターネット上のHTML(Hyper Text Markup Language)文章である。

【0003】

また、音声合成装置から上記のようなテキストデータを音声で出力する場合、殆どの場合が、1入力からのテキストデータを、音声合成装置に予め設定されている1種類の音声で出力していた。

【0004】

【発明が解決しようとする課題】

しかしながら、上記従来の音声合成装置においては、同時に複数のテキストデ

ータの入力を受け付け、しかもそれらの合成音声を聞き分け易く、且つできるだけ待ち時間も少なくなるように音声出力することができないという問題があった。

【0005】

本発明は、上述した点に鑑みなされたものであり、複数のテキストデータの合成音声の再生タイミングが重なってしまうのを検知した場合、現在再生中の音声波形の有無や再生待ちの音声波形の数に応じて、音声再生のスピードを早くすることによって、複数のテキストデータが同時に発声されて聞き取りにくくなることもなく、また、音声再生までの待ち時間が極力少ない状態で、再生された音声を聞けるようにした音声合成装置、音声合成システム、音声合成方法及び記憶媒体を提供することを第一の目的とする。

【0006】

また、本発明は、複数のテキストデータの合成音声の再生タイミングが繋がってしまうのを検知した場合、現在再生中の音声波形の後に、区切りを明確にするための予め定められた空白時間を設けることによって、複数のテキストデータが繋がってしまうことがなくなり、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようにした音声合成装置、音声合成システム、音声合成方法及び記憶媒体を提供することを第二の目的とする。

【0007】

また、本発明は、複数のテキストデータの合成音声の再生タイミングが繋がってしまうのを検知した場合、現在再生中の音声波形の後に、別の情報であることを判らせる特定の音声合成波形の再生を行うことによって、複数のテキストデータが繋がって発声された場合でも、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようにした音声合成装置、音声合成システム、音声合成方法及び記憶媒体を提供することを第三の目的とする。

【0008】

【課題を解決するための手段】

上記目的を達成するため、請求項1記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置であって、前記テキストデータの音声波形を生

成する音声波形生成手段と、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力する音声出力手段とを有することを特徴とする。

【0009】

上記目的を達成するため、請求項2記載の発明は、前記音声出力手段は、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力し、現時点で音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力することを特徴とする。

【0010】

上記目的を達成するため、請求項3記載の発明は、前記音声出力手段は、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする。

【0011】

上記目的を達成するため、請求項4記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置であって、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設ける音声出力手段とを有することを特徴とする。

【0012】

上記目的を達成するため、請求項5記載の発明は、前記空白時間は、任意に設定可能であることを特徴とする。

【0013】

上記目的を達成するため、請求項6記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置であって、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生する音声出力手段とを有することを特徴とする。

【0014】

上記目的を達成するため、請求項7記載の発明は、前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする。

【0015】

上記目的を達成するため、請求項8記載の発明は、遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話をを行うシステムに適用可能であることを特徴とする。

【0016】

上記目的を達成するため、請求項9記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置と、該音声合成装置に前記テキストデータを送信する外部装置とを具備してなる音声合成システムであって、前記音声合成装置は、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力する音声出力手段とを有することを特徴とする

【0017】

上記目的を達成するため、請求項10記載の発明は、前記音声合成装置の前記音声出力手段は、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力し、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力することを特徴とする。

【0018】

上記目的を達成するため、請求項11記載の発明は、前記音声合成装置の前記音声出力手段は、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする。

【0019】

上記目的を達成するため、請求項12記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置と、該音声合成装置に前記テキストデータを送信する外部装置とを具備してなる音声合成システムであって、前記音声合成装

置は、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設ける音声出力手段とを有することを特徴とする。

【0020】

上記目的を達成するため、請求項13記載の発明は、前記空白時間は、任意に設定可能であることを特徴とする。

【0021】

上記目的を達成するため、請求項14記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置と、該音声合成装置に前記テキストデータを送信する外部装置とを具備してなる音声合成システムであって、前記音声合成装置は、前記テキストデータの音声波形を生成する音声波形生成手段と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生する音声出力手段とを有することを特徴とする。

【0022】

上記目的を達成するため、請求項15記載の発明は、前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする。

【0023】

上記目的を達成するため、請求項16記載の発明は、遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話を行うシステムに適用可能であることを特徴とする。

【0024】

上記目的を達成するため、請求項17記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、前記テキストデータの音声波形を生成する音声波形生成工程と、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力する音声出力工程とを有することを特徴とする。

【0025】

上記目的を達成するため、請求項18記載の発明は、前記音声出力工程では、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力し、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力することを特徴とする。

【0026】

上記目的を達成するため、請求項19記載の発明は、前記音声出力工程では、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする。

【0027】

上記目的を達成するため、請求項20記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、前記テキストデータの音声波形を生成する音声波形生成工程と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設ける音声出力工程とを有することを特徴とする。

【0028】

上記目的を達成するため、請求項21記載の発明は、前記空白時間は、任意に設定可能であることを特徴とする。

【0029】

上記目的を達成するため、請求項22記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、前記テキストデータの音声波形を生成する音声波形生成工程と、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生する音声出力工程とを有することを特徴とする。

【0030】

上記目的を達成するため、請求項23記載の発明は、前記特定の音声合成波形

は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする。

【0031】

上記目的を達成するため、請求項24記載の発明は、遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話をを行うシステムに適用可能であることを特徴とする。

【0032】

上記目的を達成するため、請求項25記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、前記音声合成方法は、前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、複数の前記テキストデータの音声波形の再生タイミングの重なりを検知した場合は、音声波形の再生スピードを早くして出力するように制御する音声出力ステップとを有することを特徴とする。

【0033】

上記目的を達成するため、請求項26記載の発明は、前記音声出力ステップでは、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が一つの場合は、通常より若干早い再生スピードで出力するように制御し、現時点での音声再生中の音声波形があり且つ音声再生待ちの音声波形が二つ以上の場合は、更に早い再生スピードで出力するように制御することを特徴とする。

【0034】

上記目的を達成するため、請求項27記載の発明は、前記音声出力ステップでは、音声再生待ちの音声波形の数に応じた細かいステップで再生スピードを上げることが可能であることを特徴とする。

【0035】

上記目的を達成するため、請求項28記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、前記音声合成方法は、前記テキストデータの音声波形を生成するように制御する音声波形

生成ステップと、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め定められた空白時間を設けるように制御する音声出力ステップとを有することを特徴とする。

【0036】

上記目的を達成するため、請求項29記載の発明は、前記空白時間は、任意に設定可能であることを特徴とする。

【0037】

上記目的を達成するため、請求項30記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、前記音声合成方法は、前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、複数の前記テキストデータに係る音声波形を再生する場合、前の音声波形の再生終了後、次の音声波形の再生開始前に、予め用意された特定の音声合成波形を再生するように制御する音声出力ステップとを有することを特徴とする。

【0038】

上記目的を達成するため、請求項31記載の発明は、前記特定の音声合成波形は、前記前の音声波形と前記次の音声波形の間に挿入される明確に区切りと判る音声メッセージの音声合成波形であることを特徴とする。

【0039】

上記目的を達成するため、請求項32記載の発明は、遊園地等の各種施設で前記テキストデータを音声放送するシステム、インターネットを介して前記テキストデータで会話を行うシステムに適用可能であることを特徴とする。

【0040】

上記目的を達成するため、請求項33記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置であって、前記テキストデータを入力する入力手段と、前記テキストデータの音声波形を生成する音声波形生成手段と、前記音声波形に係る音声を出力する音声出力手段と、第1の前記テキストデータに係る音声の出力中に第2の前記テキストデータの入力による音声波形を検出した場

合に、前記第1のテキストデータに係る音声の出力を終了した後前記第2のテキストデータに係る音声を出力するよう前記音声出力手段を制御する制御手段とを備えることを特徴とする。

【0041】

上記目的を達成するため、請求項34記載の発明は、前記制御手段は前記第2のテキストデータによる音声波形の検出に応じて前記第1のテキストデータに係る音声波形の再生速度を通常速度よりも速くするよう前記音声出力手段を制御することを特徴とする。

【0042】

上記目的を達成するため、請求項35記載の発明は、前記制御手段は前記第1のテキストデータに係る音声出力の終了後、所定期間経過した後に前記第2のテキストデータに係る音声の出力を開始するよう前記音声出力手段を制御することを特徴とする。

【0043】

上記目的を達成するため、請求項36記載の発明は、前記制御手段は前記第1のテキストデータに係る音声出力の終了後、所定の音声を出力した後に前記第2のテキストデータに係る音声を出力するよう前記音声出力手段を制御することを特徴とする。

【0044】

上記目的を達成するため、請求項37記載の発明は、前記制御手段は前記第1のテキストデータに係る音声と前記第2のテキストデータに係る音声とを通常の再生速度で出力することを特徴とする。

【0045】

上記目的を達成するため、請求項38記載の発明は、前記音声波形生成手段により生成された音声波形データを記憶する記憶手段を備え、前記制御手段は前記記憶手段に記憶されている前記入力テキストデータに応じた音声波形データの数に応じて前記音声波形の再生速度を変更するよう前記音声出力手段を制御することを特徴とする。

【0046】

上記目的を達成するため、請求項39記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法であって、前記テキストデータを入力する入力工程と、前記テキストデータの音声波形を生成する音声波形生成工程と、前記音声波形に係る音声を出力する音声出力工程と、第1の前記テキストデータに係る音声の出力中に第2の前記テキストデータの入力による音声波形を検出した場合に、前記第1のテキストデータに係る音声の出力を終了した後前記第2のテキストデータに係る音声を出力するよう前記音声出力工程を制御する制御工程とを有することを特徴とする。

【0047】

上記目的を達成するため、請求項40記載の発明は、前記制御工程では前記第2のテキストデータによる音声波形の検出に応じて前記第1のテキストデータに係る音声波形の再生速度を通常速度よりも速くするよう前記音声出力工程を制御することを特徴とする。

【0048】

上記目的を達成するため、請求項41記載の発明は、前記制御工程では前記第1のテキストデータに係る音声出力の終了後、所定期間経過した後に前記第2のテキストデータに係る音声の出力を開始するよう前記音声出力工程を制御することを特徴とする。

【0049】

上記目的を達成するため、請求項42記載の発明は、前記制御工程では前記第1のテキストデータに係る音声出力の終了後、所定の音声を出力した後に前記第2のテキストデータに係る音声を出力するよう前記音声出力工程を制御することを特徴とする。

【0050】

上記目的を達成するため、請求項43記載の発明は、前記制御工程では前記第1のテキストデータに係る音声と前記第2のテキストデータに係る音声とを通常の再生速度で出力することを特徴とする。

【0051】

上記目的を達成するため、請求項44記載の発明は、前記音声波形生成工程に

より生成された音声波形データを記憶する記憶工程を有し、前記制御工程では前記記憶工程に記憶されている前記入力テキストデータに応じた音声波形データの数に応じて前記音声波形の再生速度を変更するよう前記音声出力工程を制御することを特徴とする。

【0052】

上記目的を達成するため、請求項4 5記載の発明は、テキストデータを合成音声に変換して出力する音声合成装置に適用される音声合成方法を実行するプログラムを記憶したコンピュータにより読み出し可能な記憶媒体であって、前記音声合成方法は、前記テキストデータを入力する入力ステップと、前記テキストデータの音声波形を生成するように制御する音声波形生成ステップと、前記音声波形に係る音声を出力するように制御する音声出力ステップと、第1の前記テキストデータに係る音声の出力中に第2の前記テキストデータの入力による音声波形を検出した場合に、前記第1のテキストデータに係る音声の出力を終了した後前記第2のテキストデータに係る音声を出力するよう前記音声出力ステップを制御する制御ステップとを有することを特徴とする。

【0053】

上記目的を達成するため、請求項4 6記載の発明は、前記制御ステップでは前記第2のテキストデータによる音声波形の検出に応じて前記第1のテキストデータに係る音声波形の再生速度を通常速度よりも速くするよう前記音声出力ステップを制御することを特徴とする。

【0054】

上記目的を達成するため、請求項4 7記載の発明は、前記制御ステップでは前記第1のテキストデータに係る音声出力の終了後、所定期間経過した後に前記第2のテキストデータに係る音声の出力を開始するよう前記音声出力ステップを制御することを特徴とする。

【0055】

上記目的を達成するため、請求項4 8記載の発明は、前記制御ステップでは前記第1のテキストデータに係る音声出力の終了後、所定の音声を出力した後に前記第2のテキストデータに係る音声を出力するよう前記音声出力ステップを制御

することを特徴とする。

【0056】

上記目的を達成するため、請求項49記載の発明は、前記制御ステップでは前記第1のテキストデータに係る音声と前記第2のテキストデータに係る音声とを通常の再生速度で出力するように制御することを特徴とする。

【0057】

上記目的を達成するため、請求項50記載の発明は、前記音声波形生成ステップにより生成された音声波形データを記憶するように制御する記憶ステップを有し、前記制御ステップでは前記記憶ステップに記憶されている前記入力テキストデータに応じた音声波形データの数に応じて前記音声波形の再生速度を変更するよう前記音声出力ステップを制御することを特徴とする。

【0058】

【発明の実施の形態】

以下、本発明の実施の形態を図面に基づいて詳細に説明する。

【0059】

【第1の実施の形態】

本発明の第1の実施の形態は、他のコンピュータ（サーバコンピュータ）から非同期に送られてくるテキストデータを音声出力するシステムにおいて、1つのテキストデータの音声出力が終了する前に、次のテキストデータが送られてきた時、先に音声出力中の発声スピードを早くして出力するシステムである。

【0060】

図1は本発明の第1の実施の形態に係る音声合成装置の構成例を示すブロック図である。本発明の第1の実施の形態に係る音声合成装置は、CPU101、ハードディスクコントローラ（HDC）102、プログラム113・辞書114・音声素片データ115を有するハードディスク（HD）103、キーボード104、ポインティングデバイス（PD）105、RAM106、通信回線インターフェイス（I/F）107、VRAM108、表示コントローラ109、モニタ110、サウンドカード111、スピーカ112を備えている。図中150はサーバコンピュータである。

【0061】

上記各部の構成を詳述すると、CPU101は、本装置全体の制御を行う中央演算処理装置であり、後述の図4のフローチャートに示す処理を実行する。ハードディスクコントローラ102は、ハードディスク103内のデータやプログラムの制御を行う。ハードディスク103内には、プログラム113と、音声波形生成部（後述）において入力された漢字かな混じり文を解析し読み情報を得る時に参照される漢字等の読みやアクセント情報が登録されている辞書114と、発声する文字列に従って音声の素片を接続する際に必要になる音声素片データ115が格納されている。

【0062】

キーボード104は、文字・数字・記号等の入力に用いる。ポインティングデバイス105は、プログラム開始などの指示に用いるものであり、例えばマウスやディジタイザなどから構成されている。RAM106は、プログラムやデータを格納する。通信回線インターフェイス107は、外部のサーバコンピュータ150との間でデータのやり取りを行う。本実施形態では通信形態はTCP/IP (Transmission Control Protocol/Internet Protocol) を用いている。表示コントローラ109は、VRAM108に格納された映像データを映像信号としてモニタ110に出力する制御を行う。サウンドカード111は、CPU101で生成されRAM106に格納された音声波形データをスピーカ112を通して出力している。

【0063】

図2は本発明の第1の実施の形態に係る音声合成装置のプログラムのモジュール関係を示す説明図である。本発明の第1の実施の形態に係る音声合成装置は、辞書114、音声素片データ115、メインルーチン初期化部201、音声処理初期化部202、通信データ処理部204、通信データ格納部206、表示文章データ格納部207、文章表示部208、音声波形生成部209（音声波形生成手段）、音声出力部210（音声出力手段）、初期化部203・受信部205を有する通信処理部211、音響パラメータ212、出力パラメータ213を備えている。

【0064】

上記各部の機能を詳述すると、本実施形態のシステムが起動すると、先ず、メインルーチン220のメインルーチン初期化部201にてプログラム全体の初期化が行われる。次に、通信処理部211の初期化部203にて通信部230の初期化が行われ、音声処理初期化部202にて音声部240の初期化が行われる。本実施形態では通信形態はTCP/IPを用いている。

【0065】

通信処理部211の初期化部203にて通信部230の初期化が終了すると、通信処理部211の受信部205が起動し、サーバコンピュータ150から音声合成装置に送信されるテキストデータを受信できるようになる。通信処理部211の受信部205にてこのテキストデータを受信すると、受信されたテキストデータは通信データ格納部206に格納される。

【0066】

メインルーチン初期化部201にてメインルーチン220全体の初期化が終了すると、通信データ処理部204は通信データ格納部206の監視を開始する。受信されたテキストデータが通信データ格納部206に格納されると、通信データ処理部204はそのテキストデータを読み込み、モニタ110に表示するための表示文章を格納する表示文章データ格納部207にテキストデータを格納する。

【0067】

文章表示部208は、表示文章データ格納部207にデータがあるのを検知すると、そのデータをモニタ110に表示できる形態に変換して、VRAM108上に置く。この結果、モニタ110に表示文章が表示されることになる。この際、テキストデータに何らかの処理を施して表示文章とする場合は、その処理を通信データ処理部204にて行う。

【0068】

また、通信データ処理部204は、受信されたテキストデータを音声波形生成部209に送り、音声波形生成部209にてテキストデータの音声波形の生成を行う。この際、テキストデータに何らかの処理を施して音声波形を生成する場合

は、その処理を通信データ処理部204にて行う。音声波形生成部209にて辞書114、音声素片データ115、音響パラメータ212を参照しながら、受信したテキストデータの音声波形を生成する。生成された波形は、音声出力部210に渡される。

【0069】

図3は本発明の第1の実施の形態に係る音声合成装置の音声出力部210の詳細構成を示す説明図である。本発明の第1の実施の形態に係る音声合成装置の音声出力部210は、一時蓄積部301、コントロール部302、音声再生部304を備えている。図中303は音声波形である。

【0070】

上記各部の機能を詳述すると、一時蓄積部301は、音声波形生成部209より送られてきた音声波形303を一時的に蓄積する。コントロール部302は、音声出力部210の全体をコントロールするものであり、常時、一時蓄積部301に音声波形303が送られてきたかどうかをチェックしており、音声波形303が送られてきた場合は、送られてきた順番に音声再生部304に送って音声再生を実行させる。この時、音声再生部304で現在音声再生が実行中の場合は、その再生が終了するのを待ってから、次の音声再生を開始させるようになっている。

【0071】

音声再生部304は、上記図2の出力パラメータ213から、音声出力に必要な予め設定されたパラメータ（サンプリングレート、データのビット数等のこと）に従って、音声波形303の再生を実行し、再生された音声データは上記図1のスピーカ112から出力される。尚、音声再生部304は、コントロール部302からの指示に従って、音声再生のスピードを調整できるように構成されている。

【0072】

図12は本発明の音声合成方法を実行するプログラム及び関連データが記憶媒体から装置に供給される概念例を示す説明図である。本発明の音声合成方法を実行するプログラム及び関連データは、フロッピディスクやCD-ROM等の記憶

媒体1201を装置1202に装備された記憶媒体ドライブ挿入口1203に挿入することで供給される。その後、プログラム及び関連データを記憶媒体1201から一旦ハードディスクにインストールしハードディスクからRAMにロードするか、或いはハードディスクにインストールせずに直接RAMにロードすることで、プログラム及び関連データを実行することが可能となる。

【0073】

この場合、本発明の第1～第3の実施の形態に係る音声合成装置において本発明の音声合成方法を実行するプログラムを実行する場合は、上記図12を参照して説明したような手順で音声合成装置にプログラム及び関連データを供給するか、或いは音声合成装置に予めプログラム及び関連データを格納しておくことで、プログラム実行が可能となる。

【0074】

図11は本発明の音声合成方法を実行するプログラム及び関連データを記憶した記憶媒体の記憶内容の構成例を示す説明図である。記憶媒体は、例えばボリューム情報1101、ディレクトリ情報1102、プログラム実行ファイル1103（図1や後述の図7の、プログラム113に相当）、プログラム関連データファイル1104（図1や後述の図7の、辞書114、音声素片データ115等に相当）等の記憶内容で構成される。本発明の音声合成方法を実行するプログラムは、後述の図4（第1の実施の形態）、図5・図6（第2の実施の形態）、図9・図10（第3の実施の形態）のフローチャートに基づきプログラムコード化されたものである。

【0075】

次に、上記の如く構成された本発明の第1の実施の形態に係る音声合成装置の動作を図4を参照しながら詳細に説明する。下記の処理は上記図1に示したCPU101の制御のもとに実行される。

【0076】

図4は音声合成装置の音声波形生成部209から音声波形が音声出力部210に送られてきた時に実行される音声再生スピードの調整処理に関する流れ図である。音声波形生成部209から音声波形が音声出力部210に送られてきた時、

先ずステップS401で、音声出力部210のコントロール部302は音声再生部304の動作状態を調べて、現在音声を出力中かどうかを確認する。その結果、音声を出力中でなければ、ステップS402にて、音声の再生スピードを通常のスピードに設定する。現在音声を出力中の場合は、ステップS403に進んで、コントロール部302は現在幾つの再生待ちの音声波形が一時蓄積部301にあるかを調べる。

【0077】

この結果、現在再生待ちの音声波形が1つだけ（即ち、今送られてきた音声波形だけ）の場合は、ステップS404に進んで、音声の再生スピードを予め定められた第1の値まで早くした設定値に設定する。他方、現在再生待ちの音声波形が2つ以上（即ち、今送られてきた音声波形以外にも再生待ちの音声波形がある）場合は、ステップS405に進んで、音声の再生スピードを予め定められた第1の値より、より早い値に設定された第2の値まで早くした設定値に設定する。

【0078】

この後、ステップS406に進んで、コントロール部302から音声再生部304に対して、上記ステップS402、ステップS404、ステップS405で設定された再生スピードへの設定を実行する。これによって、その時点から音声波形再生のスピードが変化するのである。

【0079】

上記図4の流れ図の処理の結果、現在音声再生中でない場合は、通常の再生スピードで再生し（その時点からの再生スピードの変更となるため、この場合は、今、音声出力部210に送られてきた音声波形303の再生スピードが通常の再生スピードとなる）、現在再生中の音声波形はあるが、再生待ちの音声波形が1つだけの場合は、少し早い再生スピードで再生し（その時点からの再生スピードの変更となるため、この場合は、今再生中の音声波形303の再生スピードが少し早くなつて再生される）、現在再生中の音声波形があり、且つ再生待ちの音声波形が2つ以上の場合は、更に早い再生スピードで再生する（その時点からの再生スピードの変更となるため、この場合は、今再生中の音声波形303の再生スピードが更に早くなつて再生される）ことになるのである。

【0080】

従って、複数の音声再生の要求が来た場合でも、音声再生の重なりが生じて聞き取りにくくなることもなく、また、音声再生までの待ち時間が極力少ない状態で、再生された音声を聞くことができるようになるのである。尚、上記ステップS405において、再生待ちの音声波形の数に応じて、もっと細かいステップで再生スピードを上げるように設定することも可能である。

【0081】

以上説明したように、本発明の第1の実施の形態に係る音声合成装置によれば、複数の音声出力が送られて来た時、音声の再生が重なって聞き取りにくくなることもなく、再生順番待ちの時間も極力少ない状態で音声再生を聞けるようになるという効果を奏する。

【0082】

本実施形態を用いれば、例えば遊園地で各場所から送られてきたテキスト情報を、サーバコンピュータを介して音声放送するシステムにおいて、送られてきた情報が一時的に重なった場合でも、重なって再生されて聞き取りにくくなることもなく、また、再生順番待ちの時間も極力少ない状態で再生音声を聞けるようになるという効果を奏する。

【0083】

また、本実施形態を用いれば、例えばインターネットで接続された複数のユーザ同士がサーバコンピュータを介してテキストデータで会話をを行うチャットシステムにおいて、サーバコンピュータから送られてきた他のユーザの発言であるテキストデータを音声出力する際に、複数のユーザからのテキストデータの音声出力が重なりそうになった場合に、音声が重なって再生されて聞き取りにくくなることもなく、また、再生順番待ちの時間も極力少ない状態で再生音声を聞けるようになるという効果を奏する。

【0084】

〔第2の実施の形態〕

本発明の第2の実施の形態は、他のコンピュータ（サーバコンピュータ）から非同期に送られてくるテキストデータを音声出力するシステムにおいて、1つの

テキストデータの音声出力が終了する前に、次のテキストデータが送られてきた時、先に音声出力中の音声の発声が終わった後、次の合成音声の発声を始める前に、予め定められた空白時間を設けるシステムである。また、前述の実施形態では1つのテキストデータの音声出力中に次の合成音声波形を検出した場合、各音声の再生速度を早くしたが、本形態では、両者の再生速度を特に早くせず、通常の再生速度で出力するものとする。

【0085】

本発明の第2の実施の形態に係る音声合成装置は、上記第1の実施の形態と同様に、CPU101、ハードディスクコントローラ(HDC)102、プログラム113・辞書114・音声素片データ115を有するハードディスク(HD)103、キーボード104、ポインティングデバイス(PD)105、RAM106、通信回線インターフェイス(I/F)107、VRAM108、表示コントローラ109、モニタ110、サウンドカード111、スピーカ112を備えている(上記図1参照)。尚、CPU101は、後述の図5、図6のフローチャートに示す処理を実行する。音声合成装置の各部の構成は上記第1の実施の形態で詳述したので説明を省略する。

【0086】

また、本発明の第2の実施の形態に係る音声合成装置のプログラムモジュールは、上記第1の実施の形態と同様に、辞書114、音声素片データ115、メインルーチン初期化部201、音声処理初期化部202、通信データ処理部204、通信データ格納部206、表示文章データ格納部207、文章表示部208、音声波形生成部209、音声出力部210、初期化部203・受信部205を有する通信処理部211、音響パラメータ212、出力パラメータ213を備えている(上記図2参照)。音声合成装置のプログラムモジュールの構成は上記第1の実施の形態で詳述したので説明を省略する。

【0087】

また、本発明の第2の実施の形態に係る音声合成装置の音声出力部210は、上記第1の実施の形態と同様に、一時蓄積部301、コントロール部302、音声再生部304を備えている(上記図3参照)。尚、音声再生部304で現在音

声再生が実行中の場合は、その再生が終了するのを待つようになっている。音声出力部210の各部の構成は上記第1の実施の形態で詳述したので説明を省略する。

【0088】

次に、上記の如く構成された本発明の第2の実施の形態に係る音声合成装置の動作を図5、図6を参照しながら詳細に説明する。下記の処理は上記図1に示したCPU101の制御のもとに実行される。

【0089】

図5は音声合成装置の音声波形生成部209から音声波形が音声出力部210に送られてきた時に実行される再生時の繋がりのチェックに関する流れ図である。音声波形が音声出力部210に送られてきた時、先ずステップS501で、音声出力部210のコントロール部302は現在幾つの再生待ちの音声波形が一時蓄積部301にあるか調べる。この結果、現在再生待ちの音声波形が1つだけ（即ち、今送られてきた音声波形だけ）の場合は、ステップS502に進む。他方、現在再生待ちの音声波形が2つ以上（即ち、今送られてきた音声波形以外にも再生待ちの音声波形がある）の場合は、ステップS505に進む。

【0090】

次にステップS502では、コントロール部302は音声再生部304の動作状態を調べて、現在音声を出力中かどうかを確認する。その結果、音声を出力中でなければ、ステップS503に進み、現在音声を出力中の場合は、ステップS505に進む。次にステップS503では、最終の音声出力の終了後、どの程度の時間が経過しているかをチェックする。その時間が、予め定められた所定の時間より短い場合は、ステップS506に進み、予め定められた所定の時間より同じか長い場合は、ステップS504に進む。

【0091】

ステップS504は、今届いた音声波形以外に再生待ちの音声はなく、且つ、現在再生中の音声もなく、更に、最後に再生された音声が終了してから予め定められた所定の時間以上の時間が経過している場合に実行されるステップで、ここでは、予め定められた所定時間の空白は設けないというフラグの設定を行って、

このフローの処理を終了する。

【0092】

ステップS505は、今届いた音声波形以外に再生待ちの音声があるか、且つ、現在再生中の音声がある場合に実行されるステップで、ここでは、予め定められた所定時間の空白を設けるというフラグの設定を行って、このフローの処理を終了する。この場合、上記所定時間は任意に設定可能である。

【0093】

ステップS506は、最後に再生された音声が終了してから予め定められた所定の時間が経過していない場合に実行されるステップで、ここでは、予め定められた所定時間までの不足時間の空白を設けるというフラグの設定と、その不足時間の設定を行って、このフローの処理を終了する。不足時間Tは、予め定められた所定時間を t_0 、最後に再生された音声が終了してからの経過時間を t_1 とした時に、

$$T = t_0 - t_1$$

で求めることができる。

【0094】

図6は実際の音声波形再生を実行する処理の流れ図である。先ずステップS601で、音声出力部210のコントロール部302は、現在再生待ちの音声波形が一時蓄積部301にあるかどうかを調べる。再生待ちの音声波形が一時蓄積部301にない場合は、上記ステップS601を繰り返して、音声波形が到着するのを待つ。ステップS602では、再生待ちの音声波形が一時蓄積部301にある場合に、上記図5の流れ図で示される所定時間の空白の有無を示すフラグの設定が済んでいるかを確認する。まだフラグの設定が終わっていない場合は、上記ステップS602を繰り返して、フラグが設定されるのを待つ。

【0095】

次にステップS603で、どのようなフラグが設定されたかを確認する。フラグが「所定時間の空白あり」に設定されている場合は、ステップS604に進んで、コントロール部302は予め定められた所定時間が経過するのを待ち、ステップS605に進む。このステップS605で、予め定められた所定時間が経過

するのを待つことにより、この間の音声再生が行われないため、所定の空白時間、即ち、無音の時間が生まれるのである。

【0096】

上記ステップS603でフラグが「不足時間あり」に設定されている時は、ステップS607に進んで、コントロール部302は不足時間が経過するのを待ち、ステップS605に進む。このステップS605で、不足時間が経過するのを待つことにより、この間の音声再生が行われないため、最後に再生された音声が終了してからの時間を加えて、所定の空白時間、即ち、無音の時間が生まれるのである。

【0097】

ステップS605は、ステップS603でフラグが所定時間の空白なしに設定されている場合と、ステップS604またはステップS607で所定時間・もしくは不足時間の経過を待った後で実行されるステップで、一時蓄積部301に蓄積されている最初の音声波形303を音声再生部304で再生開始する。この後、ステップS606で、この音声波形の再生が終了するを待って、上記ステップS601に戻る。

【0098】

このようにすることで、複数の音声再生の要求が重なって送られてきて、そのまま再生すると、音声が繋がってしまい、音声情報の区切りが判りづらくなるのに対し、音声情報の間に、明らかに区切りとわかる予め定められた空白を入れることにより、聞く人は容易に情報の区切りを区別できるようになるのである。

【0099】

以上説明したように、本発明の第2の実施の形態に係る音声合成装置によれば、複数の音声出力が送られてきた時、明らかに区切りとわかる予め定められた空白の挿入により、音声の再生が繋がってしまうことがなくなり、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0100】

本実施形態を用いれば、例えば遊園地で各場所から送られてきたテキスト情報

を、サーバコンピュータを介して音声放送するシステムにおいて、情報が一時的に重なって送られてきた結果、音声が繋がって再生されそうになった場合でも、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0101】

また、本実施形態を用いれば、例えばインターネットで接続された複数のユーザ同士がサーバコンピュータを介してテキストデータで会話をを行うチャットシステムにおいて、サーバコンピュータから送られてきた他のユーザの発言であるテキストデータを音声出力する際に、複数のユーザからのテキストデータが一時的に重なって送られてきた結果、音声が繋がって再生されそうになった場合でも、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0102】

[第3の実施の形態]

本発明の第3の実施の形態は、他のコンピュータ（サーバコンピュータ）から非同期に送られてくるテキストデータを音声出力するシステムにおいて、1つのテキストデータの音声出力が終了する前に、次のテキストデータが送られてきた時、先に音声出力中の音声の発声が終わった後、次の合成音声の発声を始める前に、例えば「次のお知らせをいたします」のような、予め用意された特定の合成音声の発声を行わせるシステムである。

【0103】

図7は本発明の第3の実施の形態に係る音声合成装置の構成例を示すブロック図である。本発明の第3の実施の形態に係る音声合成装置は、CPU101、ハードディスクコントローラ（HDC）102、プログラム113・辞書114・音声素片データ115・特定音声合成波形116を有するハードディスク（HD）103、キーボード104、ポインティングデバイス（PD）105、RAM106、通信回線インターフェイス（I/F）107、VRAM108、表示コントローラ109、モニタ110、サウンドカード111、スピーカ112を備えている。図中150はサーバコンピュータである。

【0104】

第3の実施の形態が上記第1～第2の実施の形態と相異する点を説明すると、C P U 1 0 1は、図9、図10のフローチャートに示す処理を実行する。ハードディスク103内に格納された特定音声合成波形116は、2つの音声合成が繋がりそうな場合に使用される、例えば「次のお知らせをいたします」のような、特定の音声合成波形である。音声合成装置の各部の構成は上記第1の実施の形態で詳述したので説明を省略する。

【0105】

図8は本発明の第3の実施の形態に係る音声合成装置のプログラムのモジュール関係を示す説明図である。本発明の第3の実施の形態に係る音声合成装置は、辞書114、音声素片データ115、メインルーチン初期化部201、音声処理初期化部202、通信データ処理部204、通信データ格納部206、表示文章データ格納部207、文章表示部208、音声波形生成部209、音声出力部210、初期化部203・受信部205を有する通信処理部211、音響パラメータ212、出力パラメータ213、特定音声合成波形116を備えている。音声合成装置の特定音声合成波形116以外の、プログラムモジュールの各部の構成は上記第1の実施の形態で詳述したので説明を省略する。

【0106】

また、本発明の第3の実施の形態に係る音声合成装置の音声出力部210は、上記第1の実施の形態と同様に、一時蓄積部301、コントロール部302、音声再生部304を備えている（上記図3参照）。尚、音声再生部304は、コントロール部302からの指示に従って、上記図8に示された特定音声合成波形116の再生もできるようになっている。音声出力部210の各部の構成は上記第1の実施の形態で詳述したので説明を省略する。

【0107】

次に、上記の如く構成された本発明の第3の実施の形態に係る音声合成装置の動作を図9、図10を参照しながら詳細に説明する。下記の処理は上記図1に示したC P U 1 0 1の制御のもとに実行される。

【0108】

図9は音声合成装置の音声波形生成部209から音声波形が音声出力部210に送られてきた時に実行される再生時の繋がりのチェックに関する流れ図である。音声波形が音声出力部210に送られてきた時、先ずステップS901で、音声出力部210のコントロール部302は現在幾つの再生待ちの音声波形が一時蓄積部301にあるか調べる。この結果、現在再生待ちの音声波形が1つだけ（即ち、今送られてきた音声波形だけ）の場合は、ステップS902に進む。他方、現在再生待ちの音声波形が2つ以上（即ち、今送られてきた音声波形以外にも再生待ちの音声波形がある）の場合は、ステップS905に進む。

【0109】

次にステップS902では、コントロール部302は音声再生部304の動作状態を調べて、現在音声を出力中かどうかを確認する。その結果、音声を出力中でなければ、ステップS903に進み、現在音声を出力中の場合は、ステップS905に進む。次にステップS903では、最終の音声出力の終了後、どの程度の時間が経過しているかをチェックする。その時間が、予め定められた所定の時間より短い場合は、ステップS905に進み、予め定められた所定の時間より同じか長い場合は、ステップS904に進む。

【0110】

ステップS904は、今届いた音声波形以外に再生待ちの音声はなく、且つ、現在再生中の音声もなく、更に、最後に再生された音声が終了してから予め定められた所定の時間以上の時間が経過している場合に実行されるステップで、ここでは、特定音声合成波形の再生は行わないというフラグの設定を行って、このフローの処理を終了する。ステップS905は、今届いた音声波形以外に再生待ちの音声があるか、現在再生中の音声があるか、最後に再生された音声が終了してから予め定められた所定の時間以上の時間が経過していない場合に実行されるステップで、ここでは、特定音声合成波形の再生を行うというフラグの設定を行って、このフローの処理を終了する。

【0111】

図10は実際の音声波形再生を実行する処理の流れ図である。先ずステップS1001で、音声出力部210のコントロール部302は、現在再生待ちの音声

波形が一時蓄積部301にあるかどうかを調べる。再生待ちの音声波形が一時蓄積部301にない場合は、上記ステップS1001を繰り返して、音声波形が到着するのを待つ。ステップS1002では、再生待ちの音声波形が一時蓄積部301にある場合に、上記図9の流れ図で示される特定音声合成波形の再生の有無を示すフラグの設定を確認する。まだフラグの設定が終わっていない場合は、上記ステップS1002を繰り返して、フラグが設定されるのを待つ。

【0112】

フラグが「再生あり」に設定されている場合は、ステップS1003に進んで、コントロール部302は上記図8の116に示される特定音声合成波形を読み出して、音声再生部304で再生を開始する。ステップS1004では、このステップS1003で開始した特定音声合成波形の再生が終了するのを待ち、ステップS1005に進む。

【0113】

ステップS1005は、上記ステップS1002でフラグが「再生なし」に設定されている場合と、上記ステップS1003とステップS1004で特定音声合成波形の再生が終了した後に実行されるステップで、この音声波形を音声再生部304で再生を開始する。この後、ステップS1006で、この音声波形の再生が終了するのを待って、上記ステップS1001に戻る。

【0114】

このようにすることで、複数の音声再生の要求が重なって送られてきて、そのまま再生すると、音声が繋がってしまい、音声情報の区切りが判りづらくなるのに対し、音声情報の間に、明らかに区切りとわかる、例えば「次のお知らせをいたします」のような、特定の音声合成波形の再生を入れることにより、聞く人は容易に情報の区切りを区別できるようになるのである。

【0115】

以上説明したように、本発明の第3の実施の形態に係る音声合成装置によれば、複数の音声出力が送られてきた時、音声の再生が繋がって聞き取りにくくなつても、明らかに区切りとわかる特定の音声合成波形の挿入により、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効

果を奏する。

【0116】

本実施形態を用いれば、例えば遊園地で各場所から送られてきたテキスト情報を、サーバコンピュータを介して音声放送するシステムにおいて、情報が一時的に重なって送られてきた結果、音声が繋がって再生された場合でも、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0117】

また、本実施形態を用いれば、例えばインターネットで接続された複数のユーザー同士がサーバコンピュータを介してテキストデータで会話を行うチャットシステムにおいて、サーバコンピュータから送られてきた他のユーザの発言であるテキストデータを音声出力する際に、複数のユーザからのテキストデータが一時的に重なって送られてきた結果、音声が繋がって再生された場合でも、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0118】

【他の実施の形態】

上述した本発明の実施の形態においては、音声合成装置を適用する具体例として、遊園地でテキストデータを音声放送する場合を例に上げたが、これ以外に、例えばモーターショー等の各種催事施設における催事案内／館内呼出し等に関する音声放送、カーレース場等の各種スポーツ施設におけるレース競技案内／館内呼出し等に関する音声放送など、各種分野に適用可能であり、上記実施形態と同様の効果が得られる。

【0119】

尚、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器からなる装置に適用してもよい。上述した実施形態の機能を実現するソフトウェアのプログラムコードを記憶した記憶媒体を、システム或いは装置に供給し、そのシステム或いは装置のコンピュータ（またはCPUやMPU）が記憶媒体等の媒体に格納されたプログラムコードを読み出し実行することによっても、達成さ

れることは言うまでもない。

【0120】

この場合、記憶媒体等の媒体から読み出されたプログラムコード 자체が上述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体等の媒体は本発明を構成することになる。プログラムコードを供給するための記憶媒体等の媒体としては、例えば、フロッピディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROM、或いはネットワークを介してのダウンロードなどの方法を用いることができる。

【0121】

また、コンピュータが読み出したプログラムコードを実行することにより、上述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOSなどが実際の処理の一部または全部を行い、その処理によって上述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0122】

更に、記憶媒体等の媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって上述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0123】

【発明の効果】

以上説明したように、請求項1～3、8記載の音声合成装置、請求項9～11、16記載の音声合成システム、請求項17～19、24記載の音声合成方法、請求項25～27、32記載の記憶媒体によれば、複数のテキストデータの合成音声の再生タイミングが重なってしまうのを検知した場合、現在再生中の音声波形の有無や再生待ちの音声波形の数に応じて、音声再生のスピードを早くするこ

とによって、複数のテキストデータが同時に発声されて聞き取りにくくなることもなく、また、音声再生までの待ち時間が極力少ない状態で、再生された音声を聞けるようになるという効果を奏する。

【0124】

また、請求項4～5、8記載の音声合成装置、請求項12～13、16記載の音声合成システム、請求項20～21、24記載の音声合成方法、請求項28～29、32記載の記憶媒体によれば、複数のテキストデータの合成音声の再生タイミングが繋がってしまうのを検知した場合、現在再生中の音声波形の後に、区切りを明確にするための予め定められた空白時間を設けることによって、複数のテキストデータが繋がってしまうことがなくなり、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0125】

また、請求項6～7、8記載の音声合成装置、請求項14～15、16記載の音声合成システム、請求項22～23、24記載の音声合成方法、請求項30～31、32記載の記憶媒体によれば、複数のテキストデータの合成音声の再生タイミングが繋がってしまうのを検知した場合、現在再生中の音声波形の後に、別的情報であることを判らせる特定の音声合成波形の再生を行うことによって、複数のテキストデータが繋がって発声された場合でも、音声情報の区切りが明確にわかるので、音声情報を容易に区別して聞けるようになるという効果を奏する。

【0126】

また、請求項33～38記載の音声合成装置、請求項39～44記載の音声合成方法、請求項45～50記載の記憶媒体においても、上記と同様に、複数のテキストデータが同時に発声されて聞き取りにくくなることもなく、また、音声再生までの待ち時間が極力少ない状態で、再生された音声を聞けるようになるという効果を奏する。

【図面の簡単な説明】

【図1】

本発明の第1～第2の実施の形態に係る音声合成装置の全体構成例を示すプロック図である。

【図2】

本発明の第1～第2の実施の形態に係る音声合成装置のプログラムのモジュールの構成例を示す説明図である。

【図3】

本発明の第1～第3の実施の形態に係る音声合成装置のプログラムのモジュールにおける音声出力部の詳細構成例を示す説明図である。

【図4】

本発明の第1の実施の形態に係る音声合成装置の音声波形生成部から音声波形が音声出力部に送られてきた時に実行する音声再生スピードの調整処理を示す流れ図である。

【図5】

本発明の第2の実施の形態に係る音声合成装置の音声波形生成部から音声波形が音声出力部に送られてきた時に実行する音声の繋がりのチェック処理を示す流れ図である。

【図6】

本発明の第2の実施の形態に係る音声合成装置の音声出力部で実際の音声波形再生を実行する処理を示す流れ図である。

【図7】

本発明の第3の実施の形態に係る音声合成装置の全体構成例を示すブロック図である。

【図8】

本発明の第3の実施の形態に係る音声合成装置のプログラムのモジュールの構成例を示す説明図である。

【図9】

本発明の第3の実施の形態に係る音声合成装置の音声波形生成部から音声波形が音声出力部に送られてきた時に実行する音声の繋がりのチェック処理を示す流れ図である。

【図10】

本発明の第3の実施の形態に係る音声合成装置の音声出力部で実際の音声波形

再生を実行する処理を示す流れ図である。

【図11】

本発明の音声合成方法を実行するプログラム及び関連データを記憶した記憶媒体の記憶内容の構成例を示す説明図である。

【図12】

本発明の音声合成方法を実行するプログラム及び関連データが記憶媒体から装置に供給される概念例を示す説明図である。

【符号の説明】

101 CPU

116 特定音声合成波形

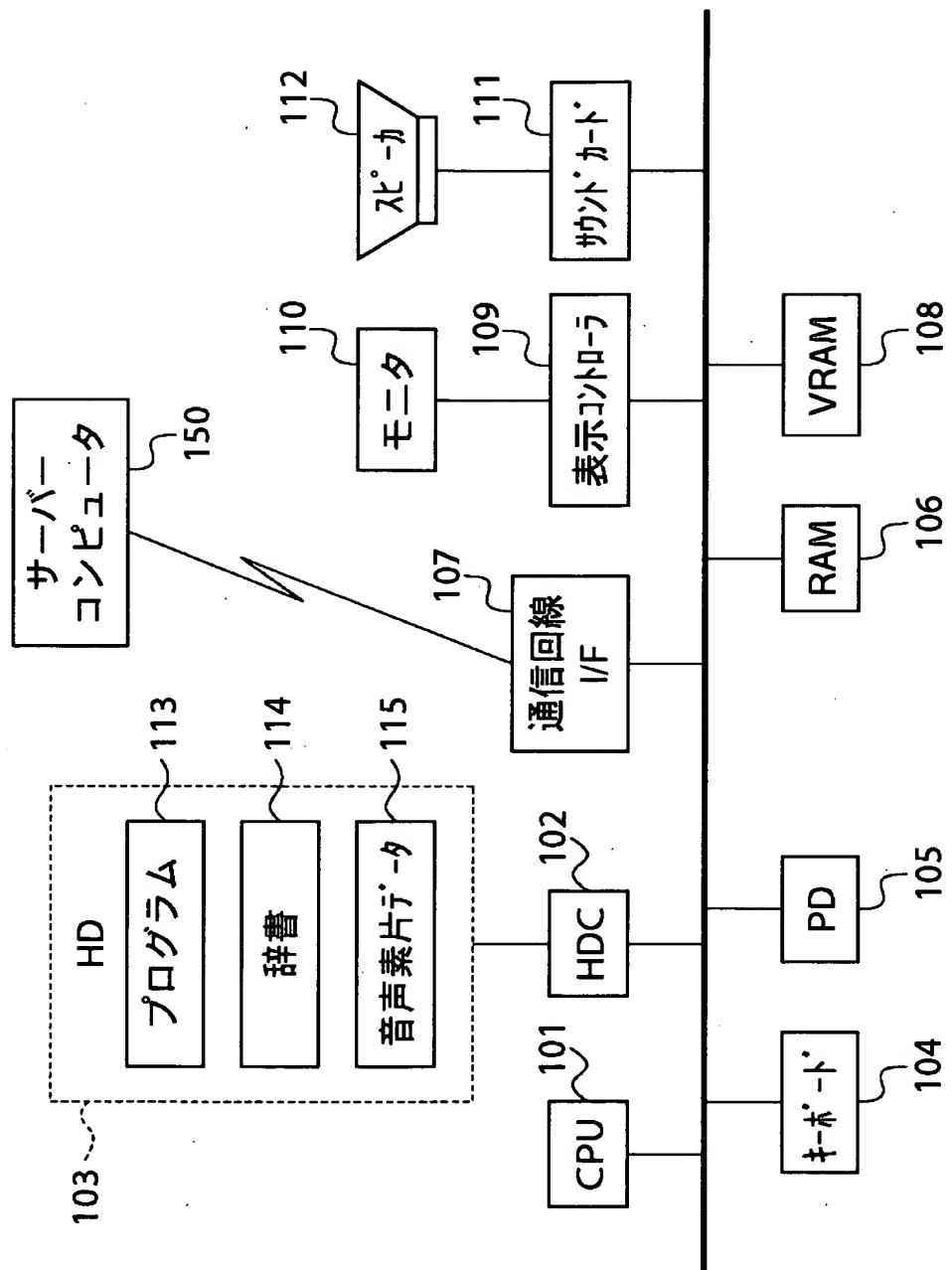
209 音声波形生成部

210 音声出力部

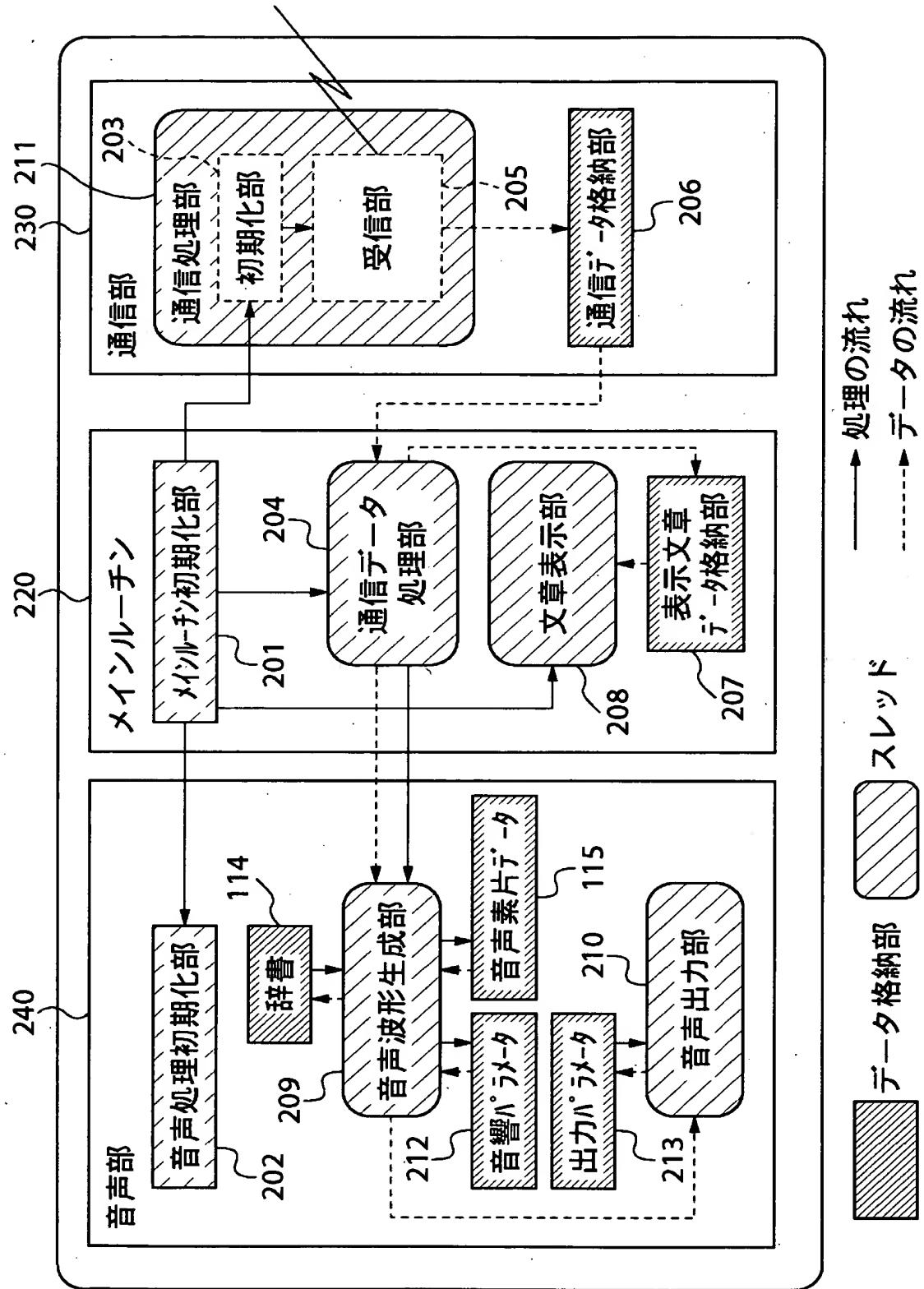
303 音声波形

【書類名】図面

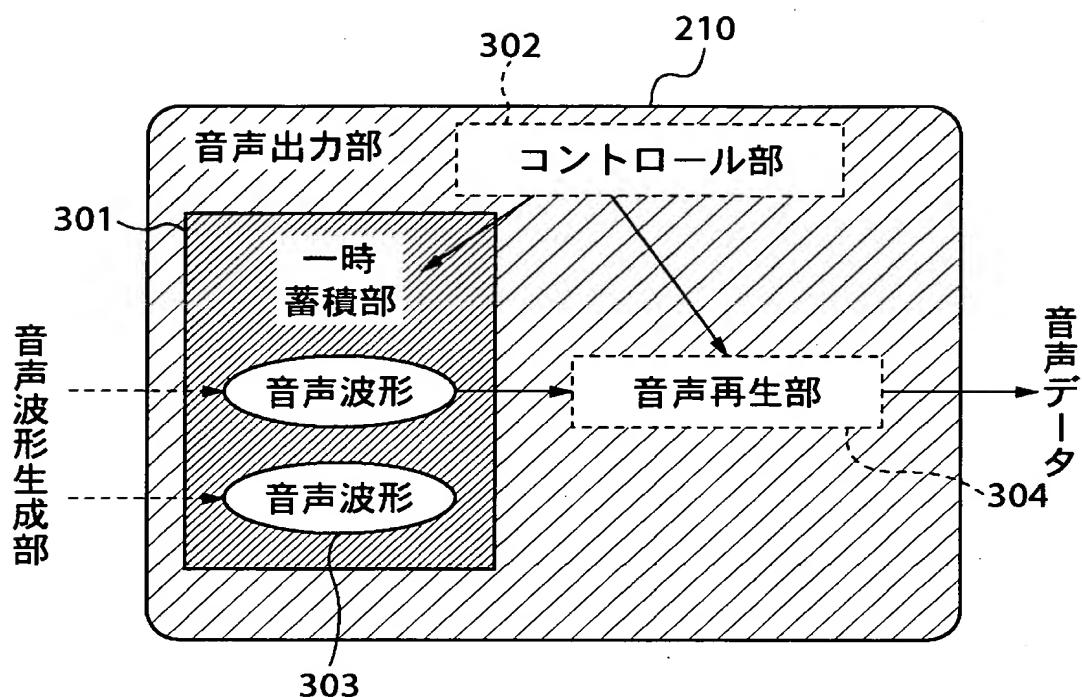
【図1】



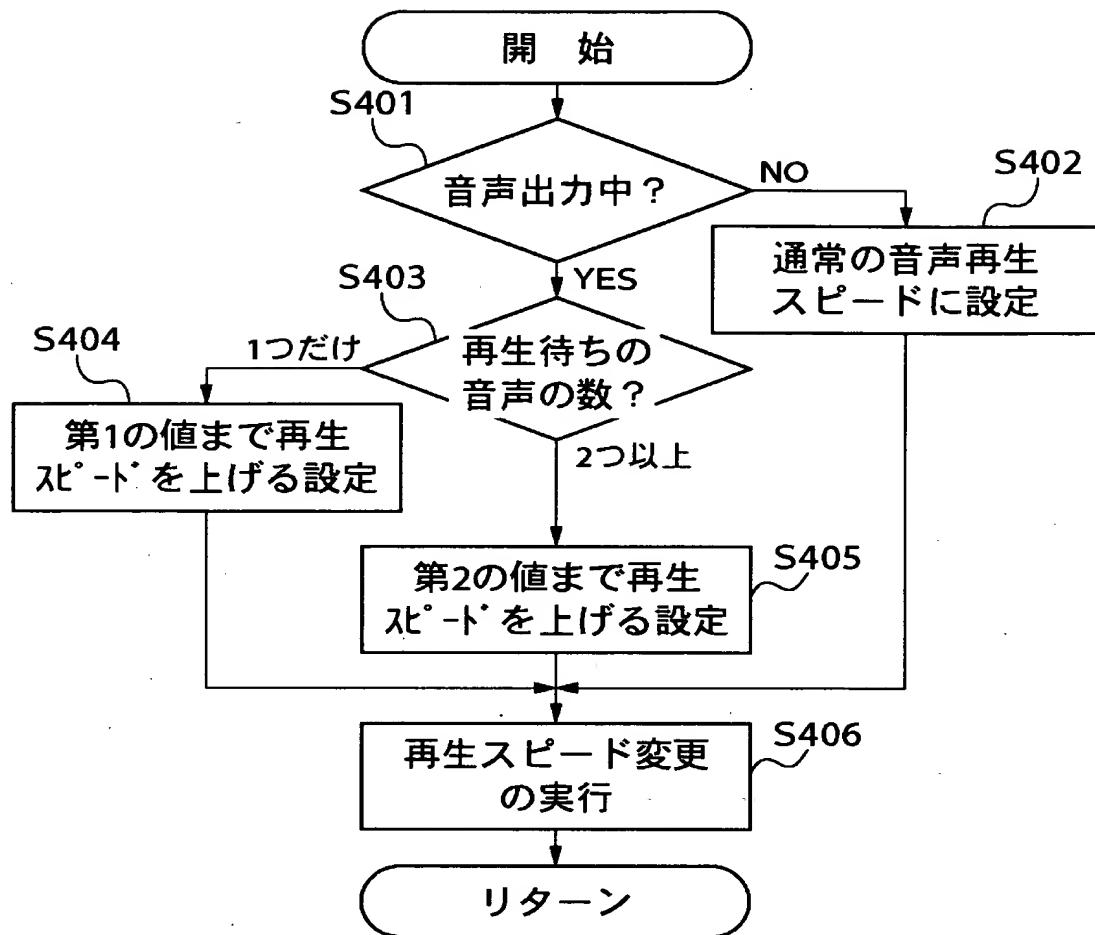
【図2】



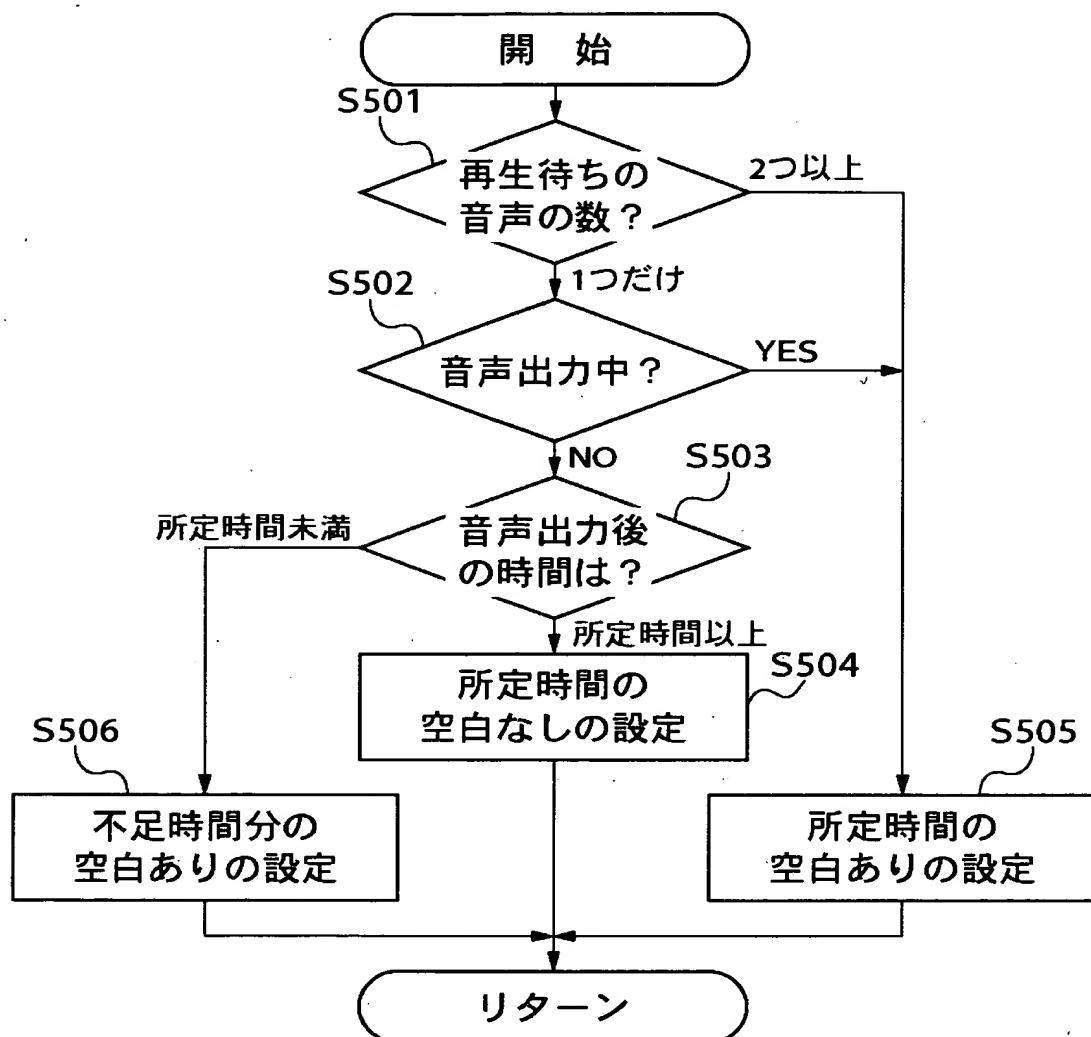
【図3】



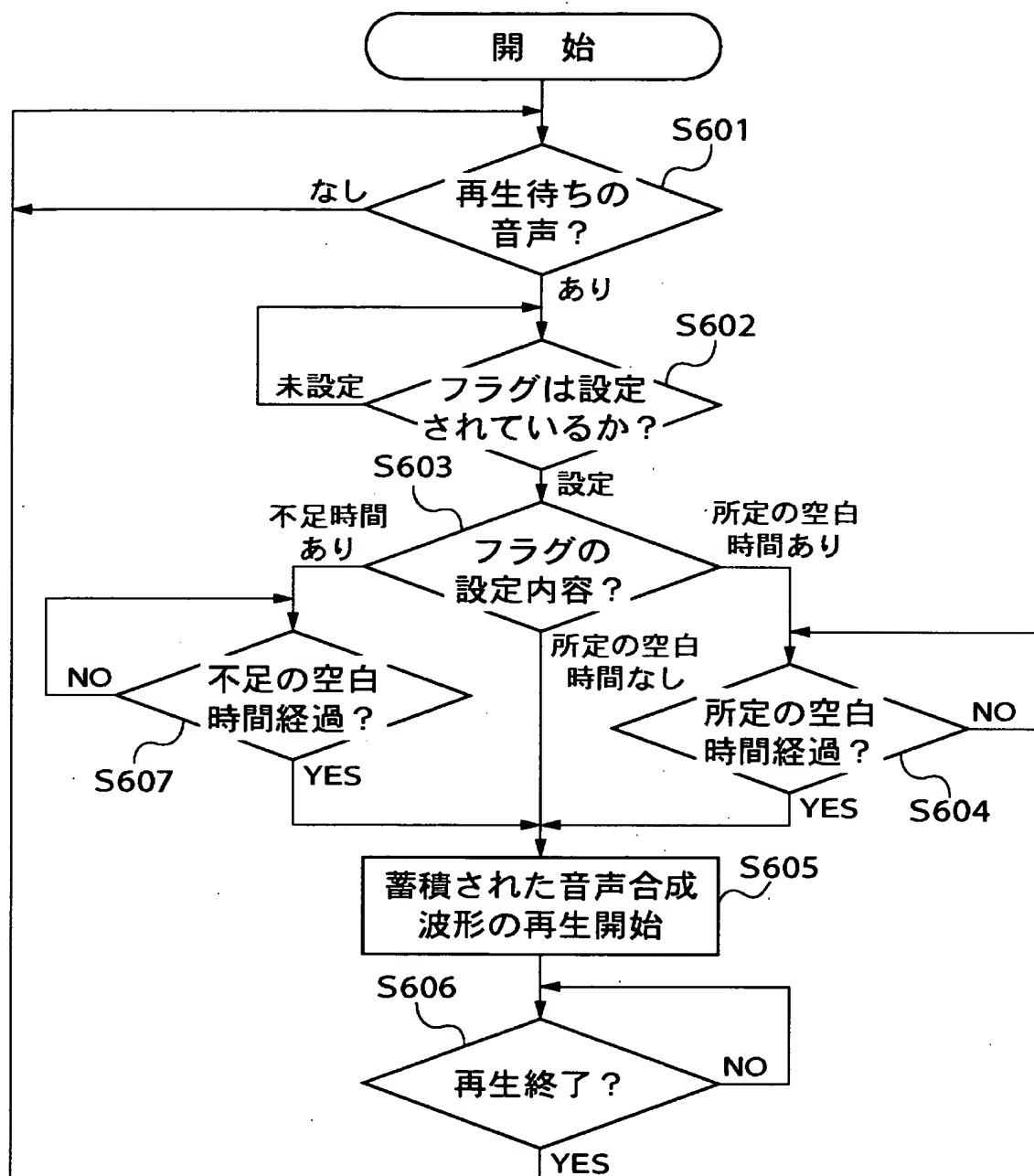
【図4】



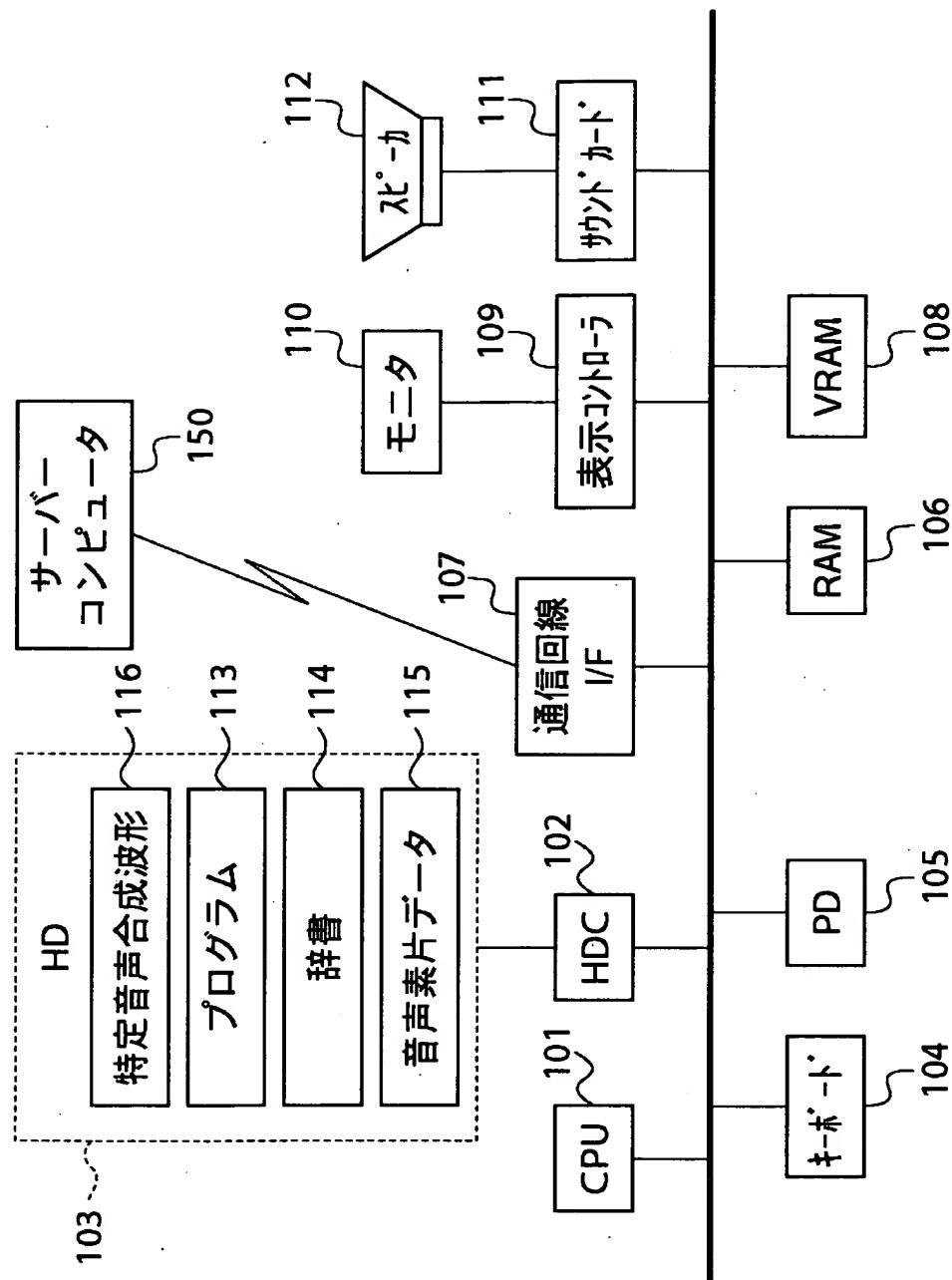
【図5】



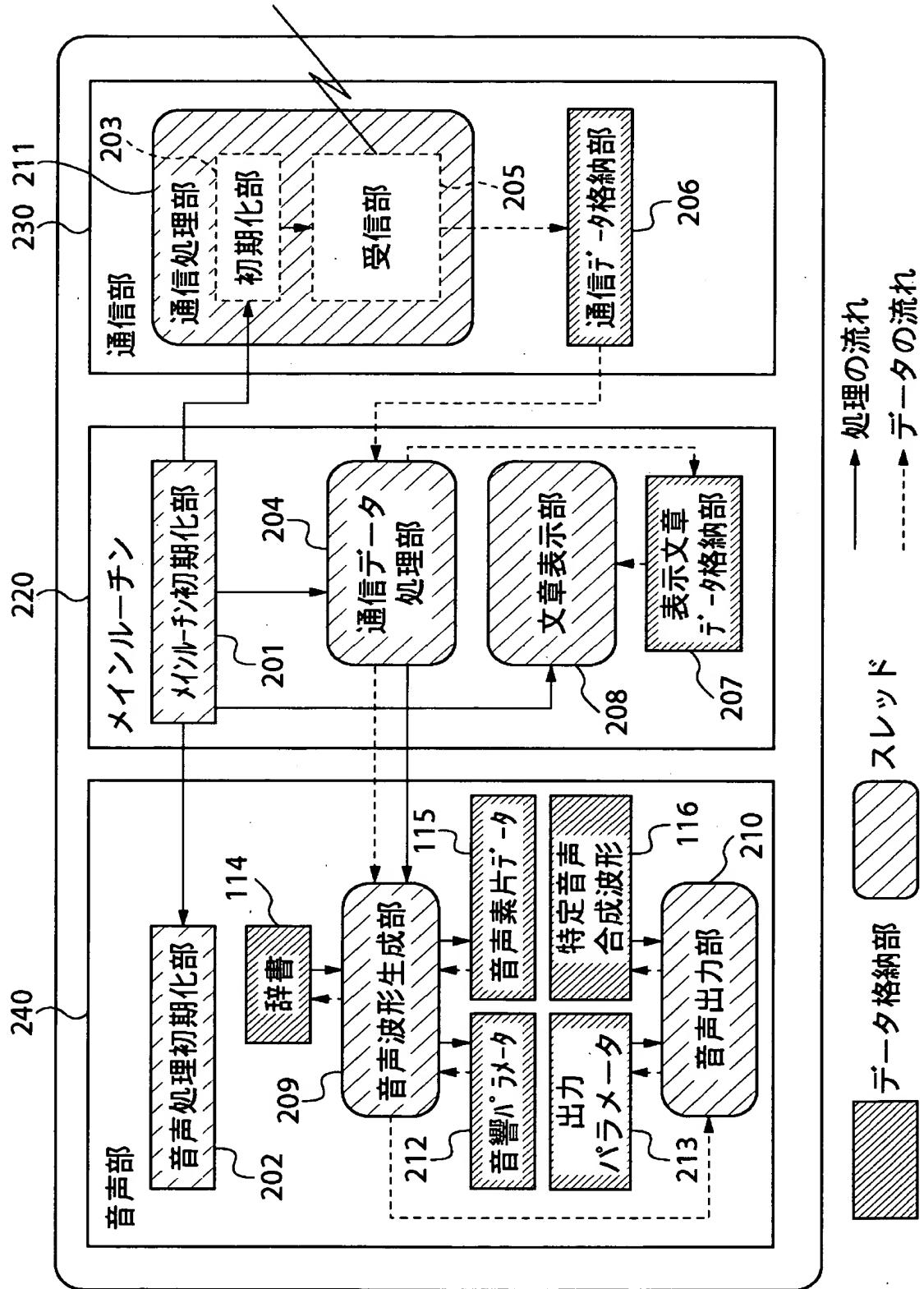
【図6】



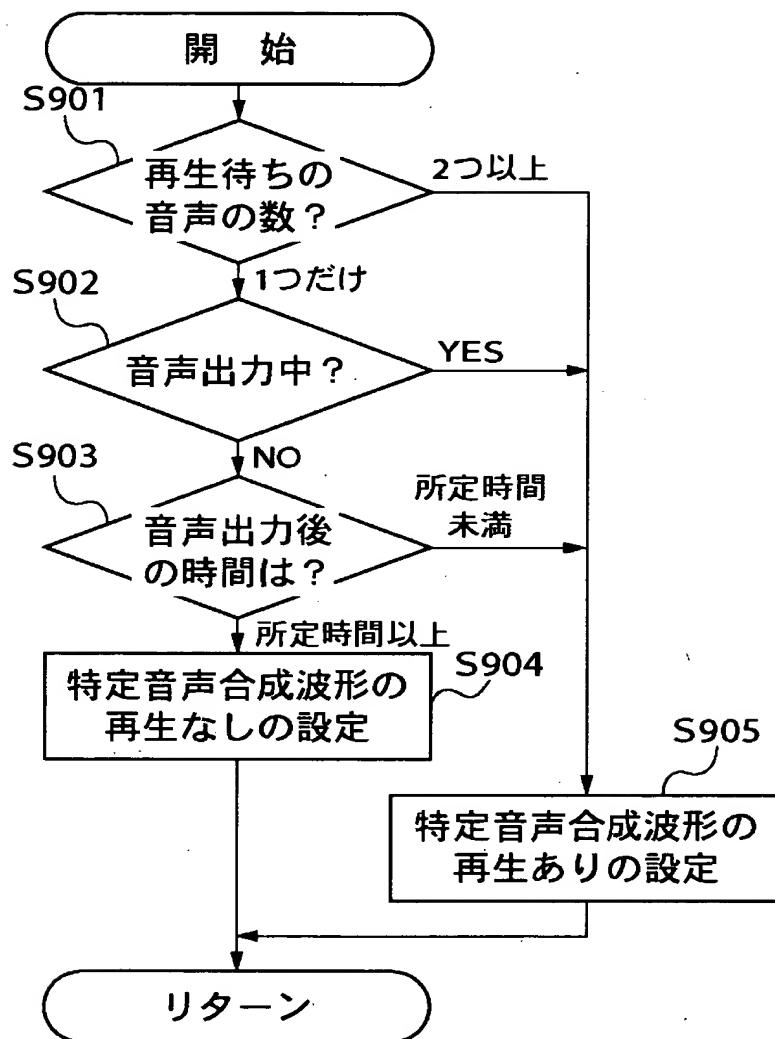
【図7】



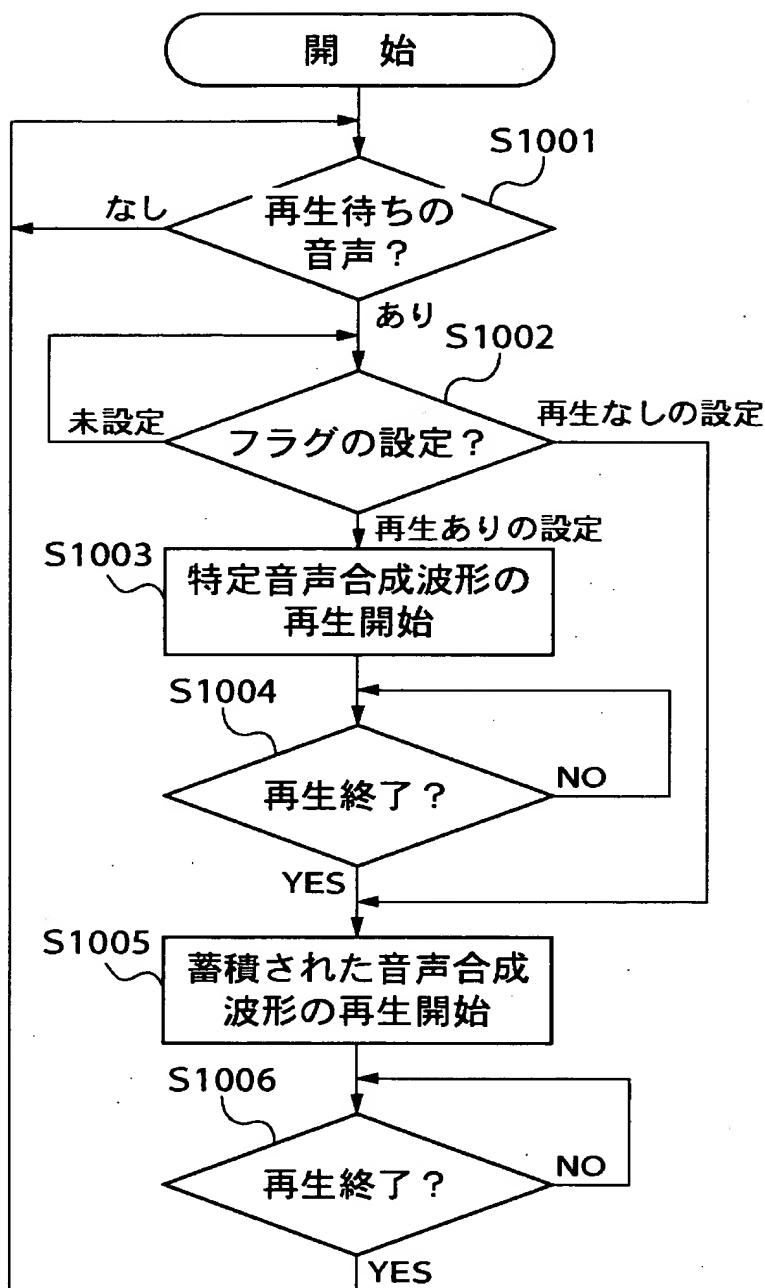
【図8】



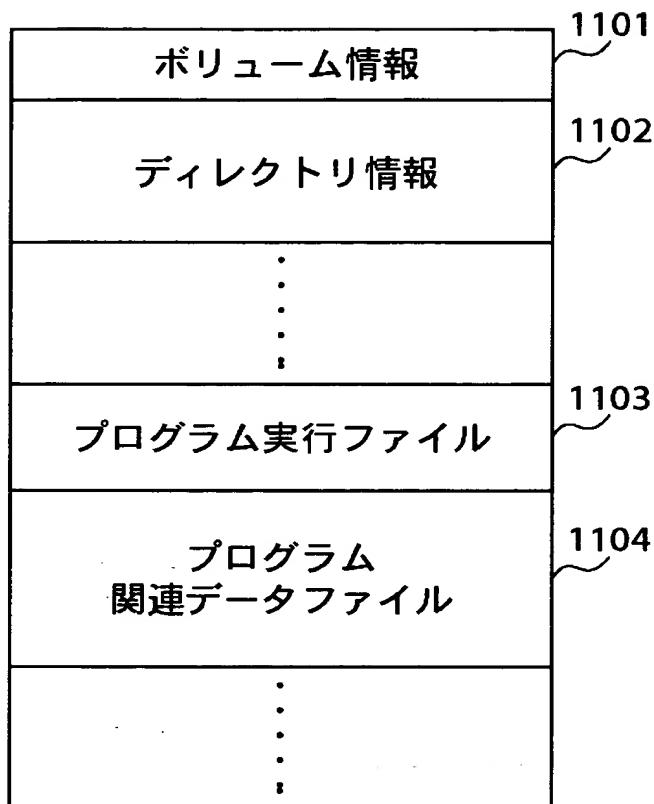
【図9】



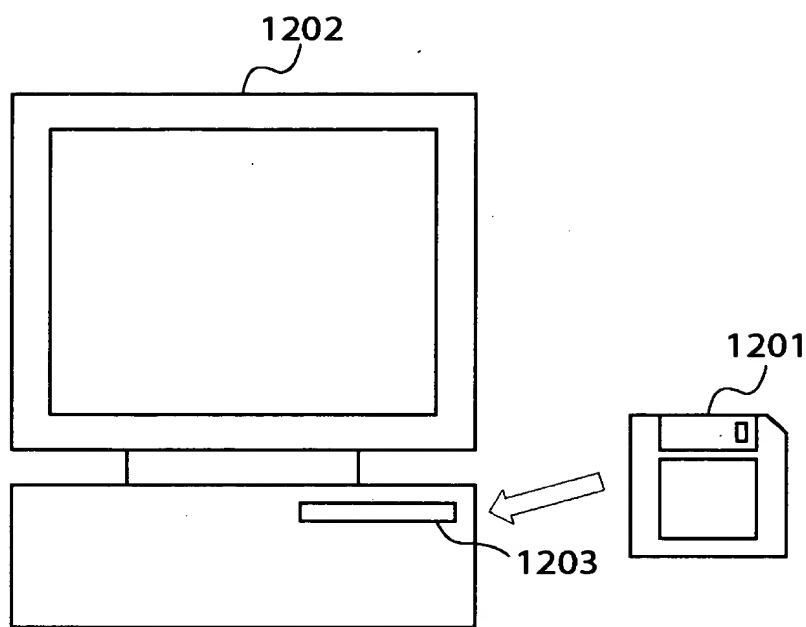
【図10】



【図11】



【図12】



【書類名】 要約書

【要約】

【課題】 複数のテキストデータの合成音声の再生タイミングが重なってしまうのを検知した場合、現在再生中の音声波形の有無や再生待ちの音声波形の数に応じて、音声再生のスピードを早くすることによって、複数のテキストデータが同時に発声されて聞き取りにくくなることもなく、また、音声再生までの待ち時間が極力少ない状態で、再生された音声を聞けるようにした音声合成装置、音声合成システム、音声合成方法及び記憶媒体を提供する。

【解決手段】 テキストデータの音声波形を生成する音声波形生成部209と、複数のテキストデータの音声波形のタイミングの重なりを検知した場合、音声波形の再生スピードを早くし、または空白を挿入し、または特定の音声合成波形を挿入して出力する音声出力部210とを具備する。

【選択図】 図2

出願人履歴情報

識別番号 [000001007]

1. 変更年月日 1990年 8月30日

[変更理由] 新規登録

住 所 東京都大田区下丸子3丁目30番2号

氏 名 キヤノン株式会社